UNSUPERVISED DEFORMABLE REGISTRATION FOR MULTI-MODAL IMAGES

RELATED APPLICATIONS

[0001] The present patent document claims the benefit of the filing date under 35 U.S.C. §119(e) of Provisional U.S. Patent Application Ser. No. 62/777,837, filed Dec. 11, 2018, which is hereby incorporated by reference.

FIELD

[0002] The present embodiments are related to image registration.

BACKGROUND

[0003] Different medical image modalities, such as magnetic resonance imaging (MRI), computed tomography (CT), and positron emission tomography (PET), show unique tissue features at different spatial resolutions. In clinical practice, multiple image modalities may be fused for diagnostic or interventional purpose, providing the combination of complementary information. Images from different modalities, however, are often acquired with different scanners and at different time points with some intra-patient anatomical changes. Multi-modal images are to be registered for an accurate analysis and interpretation.

[0004] Multi-modal image registration is a challenging problem due to the unknown and complex relationship between intensity distributions of the images to be aligned. Also, features may be present in one modality but missing in another. Previous multi-modal image approaches either rely on information theoretic measures such as mutual information or on landmarks being identified in both images. Information theoretic measures, however, often ignore spatial information, and anatomical landmarks may not always be localized in both images. Further, landmark detection may be time-consuming or may not be possible in image-guided intervention.

SUMMARY

[0005] In order to reduce computation time and provide more accurate solutions for bi-directional, multi-modal image registration, a learning-based unsupervised multi-modal deformable image registration method that does not require any aligned image pairs or anatomical landmarks is provided. A bi-directional registration function is learned based on disentangled shape representation by optimizing a similarity criterion defined on both latent space and image space.

[0006] In an embodiment, a method for unsupervised multi-modal image registration includes acquiring a first image generated by a first medical imaging modality. A second image generated by a second medical imaging modality is acquired. The second imaging modality is different than the first medical imaging modality. A prediction of deformation fields is generated between the first image and the second image. The deformation fields are generated by a machine-learned generator having been trained in domain-invariant space with machine-learned discriminators having being trained in image space.

[0007] In an embodiment, the first medical imaging modality is one of a magnetic resonance imaging (MRI) device, a computed tomography (CT) device, a positron emission tomography (PET) device, an ultrasound device, a dynaCT device, an angiogram device, and a mammography device, and the second medical imaging modality is another of the MRI device, the CT device, the PET device, the ultrasound device, the dynaCT device, the angiogram device, and the mammography device.

[0008] In an embodiment, the first image and the second image are three-dimensional (3D) images, respectively.

[0009] In an embodiment, generating includes generating with the machine-learned generator having been trained with learned shape features in the domain-invariant space decomposed from multi-modal image pairs representing a region of interest. Images of each of the multi-modal image pairs have been generated by the first medical imaging modality and the second medical imaging modality, respectively.

[0010] In an embodiment, generating includes generating with the machine-learned generator having been trained with the shape images decomposed by machine-learned encoders of an image translation network.

[0011] In an embodiment, images of the multi-modal image pairs are not aligned.

[0012] In an embodiment, generating includes generating by the machine-learned generator having been trained with latent similarity loss, and the discriminators having been trained with adversarial loss.

[0013] In an embodiment, registering includes registering the first image with the second image using a first of the deformation fields when the first image is a moving image and the second image is a fixed image, and registering the second image with the first image using a second of the deformation fields when the second image is the moving image and the first image is the fixed image.

[0014] In an embodiment, the discriminators are desriminators of a generative adversarial network (GAN).

[0015] In an embodiment, generating includes generating with the machine-learned discriminators having been trained based on translated images from a machine-learned image translation network.

[0016] In an embodiment, a system for unsupervised multi-modal image registration includes a memory configured to store a first image and a second image. The first image is generated by a first modality, and the second image is generated by a second modality. The system also includes an image processor in communication with the memory. The image processor is configured to generate a prediction of deformation fields between the first image and the second image in response to input of the first image and the second image to a machine-learned generator having been trained in domain-invariant space and machine-learned discriminators having been trained in image space. The image processor is further configured to register the first image and the second image using one of the predicted deformation fields.

[0017] In an embodiment, the system further includes a display in communication with the image processor. The display is configured to display the registered first image and second image, at least one of the deformation fields, or the registered first image and second image and the at least one deformation field.

[0018] In an embodiment, the system further includes the first modality in communication with the image processor, and the second modality in communication with the image processor.

[0019] In an embodiment, the first modality is one of an MRI device, a CT device, a PET device, an ultrasound